

## CLAIMS

What is claimed is:

- sub A1
1. A method for identifying a disease-influencing gene, the method comprising the steps of:
    - a) selecting individuals having a risk factor for a disease;
    - b) creating queries regarding the individuals' behaviors and environments;
    - c) storing the queries on a server;
    - d) providing each of the individuals with a remotely programmable apparatus having a user interface for communicating the queries and for receiving responses, and having communication means for communicating with the server through a communication network;
    - e) transmitting the queries from the server to each of the remotely programmable apparatuses;
    - f) transmitting the responses of the individuals to the queries from the remotely programmable apparatuses to the server;
    - g) creating a database of the individuals' behaviors and environments;
    - h) using data mining techniques to distinguish a group of individuals having similar behavioral and environmental profiles;
    - i) categorizing the group of individuals into at least two categories according to the individuals' disease progression;
    - j) determining the genotypes of the at least two categories of individuals;
    - k) using data mining techniques to find a gene difference between the at least two categories of individuals; and
    - l) identifying the disease-influencing gene.

- 1 2. The method of claim 1, wherein the disease-influencing  
2 gene is of the type which reduces the risk of developing  
3 the disease.
- 4
- 1 3. The method of claim 2, further comprising the step  
2 of using the disease-influencing gene to develop a  
3 drug candidate for reducing the risk of developing  
4 the disease.
- 5
- 1 4. The method of claim 2, further comprising the step  
2 of identifying a protein associated with the  
3 disease-influencing gene.
- 4
- 1 5. The method of claim 4, further comprising the  
2 step of using the protein to develop a drug  
3 candidate for reducing the risk of developing  
4 the disease.
- 5
- 1 6. The method of claim 1, wherein the disease-influencing  
2 gene is of the type which increases the risk of  
3 developing the disease.
- 4
- 1 7. The method of claim 6, further comprising the step  
2 of using the disease-influencing gene to develop a  
3 drug candidate for reducing the risk of developing  
4 the disease.
- 5
- 1 8. The method of claim 6, further comprising the step  
2 of identifying a protein associated with the  
3 disease-influencing gene.
- 4
- 1 9. The method of claim 8, further comprising the  
2 step of using the protein to develop a drug  
3 candidate for reducing the risk of developing  
4 the disease.
- 5

- 540  
A2
10. A method for identifying a disease-influencing gene, the method comprising the steps of:
- a) selecting individuals having a risk factor for a disease;
  - b) creating queries regarding the individuals' behaviors and environments;
  - c) storing the queries on a server;
  - d) providing each of the individuals with a remotely programmable apparatus having a user interface for communicating the queries and for receiving responses, and having communication means for communicating with the server through a communication network;
  - e) transmitting the queries from the server to each of the remotely programmable apparatuses;
  - f) transmitting the responses of the individuals to the queries from the remotely programmable apparatuses to the server;
  - g) creating a database of the individuals' behaviors and environments;
  - h) distinguishing a group of individuals having similar disease progressions;
  - i) using data mining techniques to categorize the group of individuals into at least two categories according to the individuals' behavioral and environmental profiles;
  - j) determining the genotypes of the at least two categories of individuals;
  - k) using data mining techniques to find a gene difference between the at least two categories of individuals; and
  - l) identifying the disease-influencing gene.

11. The method of claim 10, wherein the disease-influencing gene is of the type which reduces the risk of developing the disease.

12. The method of claim 11, further comprising the step of using the disease-influencing gene to develop a

3 drug candidate for reducing the risk of developing  
4 the disease.

5  
1 13. The method of claim 11, further comprising the step  
2 of identifying a protein associated with the  
3 disease-influencing gene.

4  
1 14. The method of claim 13, further comprising the  
2 step of using the protein to develop a drug  
3 candidate for reducing the risk of developing  
4 the disease.

5  
1 15. The method of claim 10, wherein the disease-influencing  
2 gene is of the type which increases the risk of  
3 developing the disease.

4  
1 16. The method of claim 15, further comprising the step  
2 of using the disease-influencing gene to develop a  
3 drug candidate for reducing the risk of developing  
4 the disease.

5  
1 17. The method of claim 15, further comprising the step  
2 of identifying a protein associated with the  
3 disease-influencing gene.

4  
1 18. The method of claim 17, further comprising the  
2 step of using the protein to develop a drug  
3 candidate for reducing the risk of developing  
4 the disease.

5  
1 19. A method for identifying a disease-influencing substance,  
2 the method comprising the steps of:

- 3 a) selecting individuals having a risk factor for a  
4 disease;  
5 b) creating queries regarding the individuals' behaviors  
6 and environments;

- 7 c) storing the queries on a server;  
8 d) providing each of the individuals with a remotely  
9 programmable apparatus having a user interface for  
10 communicating the queries and for receiving responses,  
11 and having communication means for communicating with  
12 the server through a communication network;  
13 e) transmitting the queries from the server to each of the  
14 remotely programmable apparatuses;  
15 f) transmitting the responses of the individuals to the  
16 queries from the remotely programmable apparatuses to  
17 the server;  
18 g) creating a database of the individuals' behaviors and  
19 environments;  
20 h) determining the genotypes of the individuals;  
21 i) distinguishing a group of the individuals having similar  
22 genotypes;  
23 j) categorizing the group of individuals into at least two  
24 categories according to their disease progressions; and  
25 k) using data mining techniques to find a disease-  
26 influencing substance from the behavioral and  
27 environmental profiles between the at least two classes  
28 of individuals.

29  
1 20. The method of claim 19, wherein the disease-influencing  
2 substance is of the type which reduces the risk of  
3 developing the disease.

4  
1 21. The method of claim 20, further comprising the step  
2 of using the disease-influencing substance to  
3 develop a drug candidate for reducing the risk of  
4 developing the disease.

5  
1 22. The method of claim 19, wherein the disease-influencing  
2 substance is of the type which increases the risk of  
3 developing the disease.  
4

1 23. The method of claim 22, further comprising the step  
2 of using the disease-influencing substance to  
3 develop a drug candidate for reducing the risk of  
4 developing the disease.

5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
SUB  
A4  
24. A database and data processing system for finding a  
disease-influencing gene among individuals having a risk  
factor for a disease, the database and data processing  
system comprising:

- a) a server for storing queries regarding the individuals' behavior and environment and for storing the individuals' responses to the queries;
- b) at least one remotely programmable apparatus in communication with the server, wherein the remotely programmable apparatus comprises:
  - i) a user interface for communicating the queries to the individuals and for receiving the responses; and
  - ii) communication means for receiving the queries from the server and for transmitting the responses to the server;
- c) genotyping means in communication with the server for obtaining the genotype of the individual; and
- d) data mining means in communication with the server, wherein the data mining means includes:
  - i) means for analyzing the responses in order to group the individuals having a similar behavioral and environmental profile, a similar disease progression, and a similar genotype;
  - ii) means for analyzing the responses in order to group the individuals having a similar disease progression;
  - iii) means for analyzing the responses in order to group the individuals having a similar genotype; and
  - iv) means for identifying the disease-influencing gene.

1 25. The system of claim 24, further comprising at least one  
2 monitoring device for producing measurements of a  
3 physiological condition of the individuals and for  
4 transmitting the measurements to the remotely  
5 programmable apparatus, wherein the apparatus further  
6 includes device interface means for receiving the  
7 measurements from the monitoring device and means for  
8 transmitting the measurements to the server.

9  
1 26. The system of claim 24, further comprising means for  
2 identifying a protein associated with the disease-  
3 influencing gene.

4  
1 27. A database and data processing system for use in finding a  
2 disease-influencing substance among individuals having a  
3 risk factor for a disease, the database and data processing  
4 system comprising:  
5 a) a server for storing queries regarding the individuals'  
6 behavior and environment and for storing the  
7 individuals' responses to the queries;  
8 b) at least one remotely programmable apparatus in  
9 communication with the server, wherein the remotely  
10 programmable apparatus comprises:  
11 i) a user interface for communicating the queries to  
12 the individuals and for receiving the responses; and  
13 ii) communication means for receiving the queries from  
14 the server and for transmitting the responses to the  
15 server;  
16 c) genotyping means in communication with the server for  
17 obtaining the genotype of the individual; and  
18 d) data mining means in communication with the server,  
19 wherein the data mining means includes:  
20 i) means for analyzing the responses in order to  
21 group the individuals having a similar behavioral

- 22 and environmental profile, a similar disease  
23 progression, and a similar genotype;  
24 ii) means for analyzing the responses in order to  
25 group the individuals having a similar disease  
26 progression;  
27 iii) means for analyzing the responses in order to  
28 group the individuals having a similar genotype;  
29 and  
30 iv) means for identifying the disease-influencing  
31 substance.

32  
1 28. The system of claim 27, further comprising at least one  
2 monitoring device for producing measurements of a  
3 physiological condition of the individuals and for  
4 transmitting the measurements to the remotely  
5 programmable apparatus, wherein the apparatus further  
6 includes device interface means for receiving the  
7 measurements from the monitoring device and means for  
8 transmitting the measurements to the server.

9  
1 29. The system of claim 27, further comprising means for  
2 identifying a protein associated with the disease-  
3 influencing gene.  
4

added  
C1